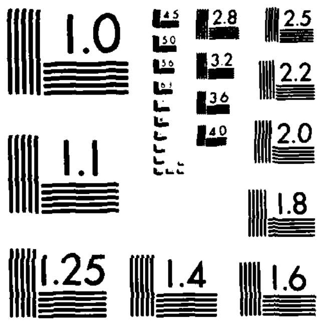
AD-A149 223 SPEECH ENHANCEMENT USING MULTIPLE MICROPHONES(U) 1/1
MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB
W A HARRISON 15 NOV 84 TR-691 ESD-TR-84-040
UNCLASSIFIED F19628-85-C-0002 F/G 17/2 NL

END
FILMED
DTIC




MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

12

Technical Report [illegible]

# Speech Enhancement Using Multiple Microphones

AD-A149 223

[illegible]

[illegible]

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

*Lexington, Massachusetts*

DTIC FILE COPY

Prepared for the Department of the Air Force
under Electronic Systems Division Contract F19628-[illegible].

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the support of the Department of the Air Force under Contract F19628-85-C-0002.







This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Thomas J. Alpert

Thomas J. Alpert, Major, USAF
Chief, ESD Lincoln Laboratory Project Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

# SPEECH ENHANCEMENT USING MULTIPLE MICROPHONES

*W.A. HARRISON*

*Group 24*

TECHNICAL REPORT 691

15 NOVEMBER 1984



LEXINGTON MASSACHUSETTS

## ABSTRACT*

A new application of Widrow's Adaptive Noise Cancelling (ANC) algorithm is presented. Specifically, the ambient environment is generalised to include the case where an acoustic barrier exists between the primary and reference microphones. By updating the coefficients of the noise estimation filter only during silence, it is shown that the ANC technique can provide substantial noise reduction with little speech distortion even when the acoustic barrier provides only moderate attenuation of acoustic signals. The use of the modified ANC method is evaluated using an oxygen facemask worn by fighter aircraft pilots. Experiments demonstrate that if a noise field is created using a single source, 11 dB signal-to-noise ratio improvement can be achieved by attaching the reference microphone to the exterior of the facemask. The length of the ANC filter required for this particular environment is only about 50 points long.



* *This report is based on a thesis of the same title submitted to the Department of* Electrical Engineering and Computer Science at the Massachusetts Institute of Technology in June 1984 in partial fulfillment for the degrees of Master of Science and Electrical Engineer.

## TABLE OF CONTENTS

## ILLUSTRATIONS

# CHAPTER 1

# INTRODUCTION

## 1.1 Problem Statement

Separating a desired signal from an undesired signal is an important and common problem in signal processing. In many military environments such as fighter aircraft, helicopters and tanks, the increasing use of digital communication systems has created a need for robust narrowband vocoders whose performance is acceptable under a variety of stressful conditions. It is known that the presence of high levels of acoustic noise in the audio signal leads to substantial reductions in the intelligibility of the synthesized speech. Although the use of gradient (noise-cancelling) microphones and oxygen facemasks reduces the noise problem somewhat, the ambient noise levels encountered in environments such as high performance tactical aircraft are often severe enough that the performance of the vocoder is seriously affected. Another application for which noise reduction is necessary is automatic speech recognition. As military hardware becomes more sophisticated, it is becoming increasingly important to allow operational personnel to perform simple tasks by voice command. Reliable recognition of noise corrupted speech has proved to be a difficult problem and the existence of noise reduction would be greatly beneficial.

Unfortunately, single channel algorithms to date [1] have been only moderately successful in reducing the unwanted signal. A potentially useful method that requires

more than one input source is Adaptive Noise Cancellation (ANC) proposed by Widrow *et al* [2]. The ANC algorithm uses a second sensor, commonly called the reference, to gain additional information about the undesired signal. If the reference sensor contains only a correlated version of the noise component in the primary sensor, one can in theory derive a linear filter which, when applied to the noise corrupted signal in the primary, minimizes the noise power while leaving the desired signal unaffected. If the noise in the reference microphone is related to the noise in the primary microphone by a linear filter, then it is possible to cancel completely the noise corrupting the desired signal. Adaptive Noise Cancelling has been applied in areas where a reference signal which contains only a correlated version of the noise in the primary input is easily obtained. Examples of such situations are echo cancellation for long distance telephone calls [3], adaptive line enhancement [4], and adaptive antenna array processing [5].

In order to apply the ANC algorithm effectively it is necessary that the reference signal contain as little of the desired signal as possible so that the algorithm does not cancel the speech along with the noise. Previous researchers [6] have met this condition by placing the reference microphone next to the noise source and placing the primary microphone close to a weak signal source located far from the reference microphone. This report presents the results of a study in which the requirement was satisfied by placing an acoustic barrier between the primary and reference microphones such that the speech component in the reference was significantly attenuated. This case is of interest since it is not always possible or feasible to place the primary and reference

microphones far apart. In practice, noise reduction techniques may be desirable in applications where the acoustic barrier can provide no more than a 15 dB difference between the signal-to-noise ratios at the primary and reference microphones. In these environments, direct application of the ANC algorithm can produce significant distortion in the processed speech due to the presence of a strong speech component at the reference microphone. Chapter 3 describes a modification of the ANC algorithm for use in such situations which involves allowing the noise cancelling filter to adapt only during silent intervals in the speech. With this modification it is possible to achieve considerable noise reduction without seriously distorting the processed speech.

A practical consideration when using the ANC algorithm is the size of the adaptive filter needed for good noise cancellation. A separation of a few meters between the two microphones requires the adaptive filter either to be able to estimate the delay between the reference and primary signals or to have a long impulse response in order to provide good cancellation of the noise in the primary channel. Boll and Pulsipher [6] needed a tapped delay line of 1500 taps to cancel the noise in the primary microphone adequately. Longer filter lengths require more computation in both estimating the filter coefficients and in filtering the signals. Also, it has been found [6] that long filter lengths tend to introduce reverberation in the processed speech due to the feedback of the speech through the adaptive filter. With an acoustic barrier present between the primary and reference microphones it is possible to place the two microphones in close proximity on either side of the barrier. The small separation between the two

microphones significantly shortens the filter length required for noise cancellation and minimizes the presence of reverberation.

The performance of the modified algorithm was evaluated using an oxygen facemask commonly worn by military fighter aircraft pilots. Since the facemask provides only moderate attenuation in an otherwise noisy cockpit environment it appeared to be an ideal candidate for evaluating the noise reduction methods described in this report. As will be discussed in subsequent chapters, experiments performed in a sound chamber demonstrated that the modified ANC algorithm can be used to provide noise cancellation with significantly less distortion introduced into the processed speech thar does the direct approach. Furthermore, the facemask permits close placement .ne two microphones and requires a filter of only 50 taps. The short filter length reduce the reverberation commonly heard in the standard ANC algorithm.

### 1.2 Report Organisation

The remainder of the report is organized as follows: Chapter 2 contains a brief discussion of the adaptive noise cancelling algorithm used in this study along with a summary of the least mean squares (LMS) and the recursive least squares (RLS) algorithms. In Chapter 3, new applications of ANC are presented. In Chapter 4, an application of ANC to the fighter cockpit environment is discussed. In Chapter 5, experimental results are presented and in Chapter 6, conclusions and possible extensions of the results are given.

# CHAPTER 2

# SUMMARY OF ADAPTIVE NOISE CANCELLATION

## 2.1 Introduction

This chapter covers the basic ideas behind adaptive noise cancellation. Also covered are two of the more common adaptive filter algorithms used in the implementation of the Adaptive Noise Cancelling (ANC) method. The ANC method and the Least Mean Squares algorithm were first proposed by Widrow and Hoff [7]. The ANC algorithm assumes there are two inputs available, a primary input which contains the desired speech plus some interfering noise and a reference input which contains only a correlated version of the interfering noise in the primary and no speech component. As will be shown in the next section, having speech in the reference sensor may cause the adaptive filter to cancel the speech signal in the primary instead of the noise. A solution to this problem is discussed in Chapter 3.

Two common adaptive filters are covered in the remaining sections of Chapter 2. These are the Least Mean Squares (LMS) and the Recursive Least Squares (RLS) algorithms. The former is a gradient based algorithm while the latter directly minimizes the sum of the square of the output signal and is commonly placed under the category of exact least squares methods.

## 2.2 Summary of Adaptive Noise Cancellation

A block diagram of the Adaptive Noise Cancelling (ANC) system is shown in Figure

2.1. For simplicity, the discrete time index will be deleted from subsequent equations. The signal $p$ at the primary microphone is composed of a desired component $s$ and a noise component $w'$. The signal $r$ at the reference microphone consists of $w$, the noise signal. The filter $H_1(\omega)$ models the relationship between the noise in the reference and primary inputs due to the separation of the two microphones. The adaptive filter $\hat{H}(\omega)$ provides an estimate $y$ of the noise in the primary channel such that $z = p - y$ is an estimate of the desired speech signal $s$. The results of Widrow *et al* [2] show that minimizing $E[z^2]$, the energy of the output signal, is equivalent to making $y$ be a least squares estimate of $w'$.

The assumptions needed for the above results are that $s$ and $w$ are wide sense stationary, zero mean processes, and that $s$ is uncorrelated with $w$ and $w'$. The output $z$ is given by

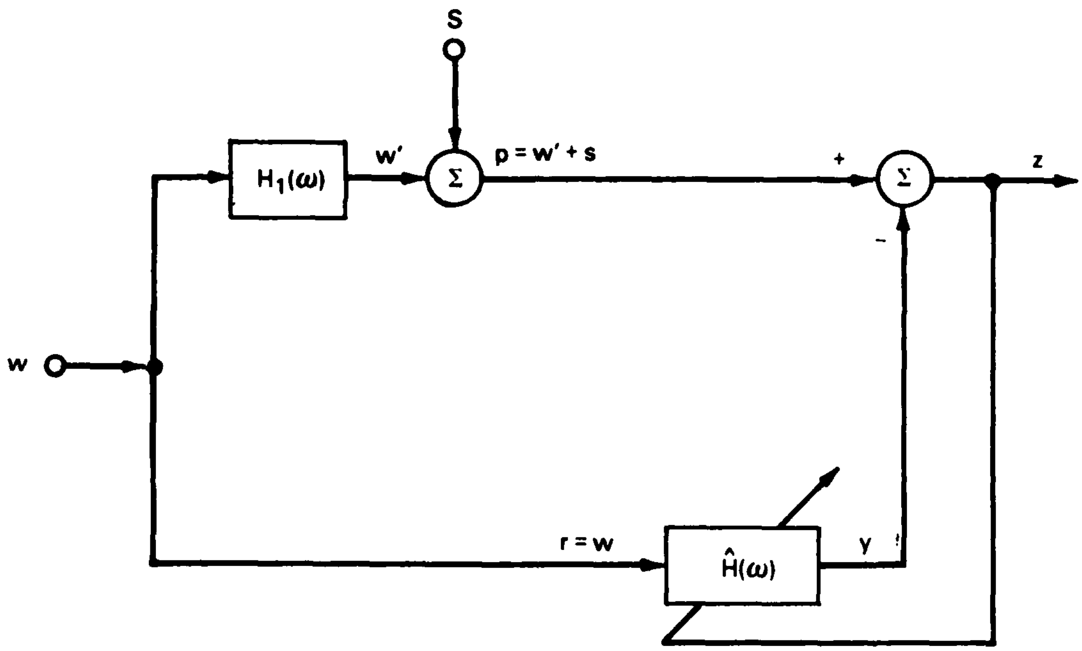
$$z = s + w' - y \tag{2.2.1}$$

Squaring both sides of equation (2.2.1) gives

$$z^2 = s^2 + 2s(w' - y) + (w' - y)^2 \tag{2.2.2}$$

Taking the expectation of both sides of (2.2.2) and noting that $s$ is uncorrelated with $w' - y$ yields

$$\begin{aligned} E[z^2] &= E[s^2] + 2E[s(w' - y)] + E[(w' - y)^2] \\ &= E[s^2] + E[(w' - y)^2] \end{aligned} \tag{2.2.3}$$




146553-N

**Figure 2.1. Block diagram of adaptive noise cancellation**

Adjusting the filter's impulse response to minimise $E[z^2]$ will not affect the signal energy $E[s^2]$. Hence the minimum output energy is

$$\min E[z^2] = E[s^2] + \min E[(w' - y)^2] \tag{2.2.4}$$

Since minimizing $E[z^2]$ corresponds to minimizing $E[(w' - y)^2]$, $y$ is the best least squares estimate of $w'$. Also, by rearranging equation (2.2.1) as

$$z - s = w' - y \tag{2.2.5}$$

one sees that the output $z$ is a best least squares estimate of the desired signal $s$.

Many methods exist to derive an $\hat{H}(\omega)$ such that the output energy is a minimum [8–12]. We will summarize two representative methods which use an FIR filter. Both recursively estimate the filter coefficients on a sample by sample basis. The Least Mean Squares (LMS) algorithm [8] is an iterative, minimum seeking method for determining the least squares solution. As will be shown in the next section, the LMS algorithm converges to the minimum of $E[z^2]$ on average. The Recursive Least Squares (RLS) algorithm [9] solves the covariance normal equations at time $n+1$ based on the solution of the covariance equation at time $n$. An advantage of the RLS algorithm is that at each time instance, the filter estimate $\hat{H}(\omega)$ exactly minimizes the error criterion $\sum_n z_n^2$. For large $n$ in a stationary, zero-mean environment, minimizing $\sum_n z_n^2$ is equivalent to minimizing $E[z^2]$.

## 2.3 Summary of the Least Mean Square Algorithm

The first adaptive filter algorithm used in this report is the Least Mean Squares (LMS)

algorithm first proposed by Widrow and Hoff [7]. The LMS algorithm is a steepest descent algorithm that uses estimates of the gradient of the error signal. To see under what conditions we expect the LMS method to converge to the global solution, let us consider the following.

We constrain our adaptive filter to be an FIR filter implemented with a tapped delay line. The results obtained can be easily extended to an FIR filter with a lattice structure [12]. A tapped delay line structure for the adaptive filter is shown in Figure 2.2. The output of the filter $y_n$, at time $n$ is given by

$$\begin{aligned} y_n &= \sum_{i=0}^{L-1} a_i r_{n-i} \\ &= \mathbf{a}_n^T \mathbf{r}_n \end{aligned} \tag{2.3.1}$$

Where

$$\mathbf{a}_n = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{L-1} \end{pmatrix} \qquad \mathbf{r}_n = \begin{pmatrix} r_n \\ r_{n-1} \\ \vdots \\ r_{n-(L-1)} \end{pmatrix}$$

The output of the system $z_n$, is

$$\begin{aligned} z_n &= p_n - y_n \\ &= p_n - \mathbf{a}_n^T \mathbf{r}_n \end{aligned} \tag{2.3.2}$$

Using the minimum mean square error as our criterion for optimality, we see that

$$\begin{aligned} E[z_n^2] &= E[(p_n - \mathbf{a}_n^T \mathbf{r}_n)^2] \\ &= E[p_n^2] - 2E[p_n \mathbf{r}_n^T]\mathbf{a}_n + \mathbf{a}_n^T E[\mathbf{r}_n \mathbf{r}_n^T]\mathbf{a}_n \end{aligned} \tag{2.3.3}$$

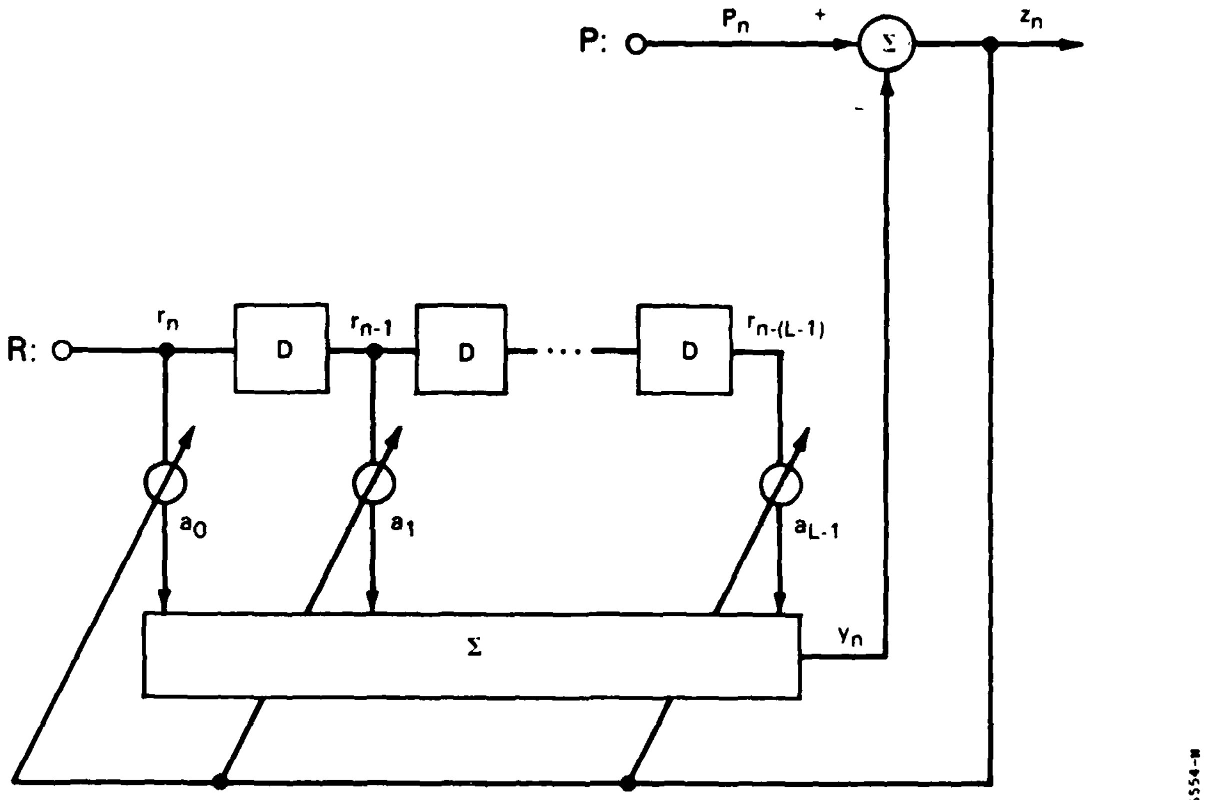



**Figure 2.2. ANC block diagram with a tapped-delay line filter**

For simplicity, define $\zeta_n = E[z_n^2]$, $\mathbf{p}_n = E[p_n \mathbf{r}_n^T]$, and $\mathbf{R}_n = E[\mathbf{r}_n \mathbf{r}_n^T]$, then equation (2.3.3) can be written as

$$\zeta_n = E\left[p_n^2\right] - 2\mathbf{p}_n \mathbf{a}_n + \mathbf{a}_n^T \mathbf{R}_n \mathbf{a}_n \tag{2.3.4}$$

What is important to note is that $\zeta_n$ is a quadratic function of $\mathbf{a}_n$ and that there exists one global minimum. The minimum mean square error $\zeta_{min}$, can be found via a steepest descent method.

If one has knowledge of $\mathbf{p}_n$ and $\mathbf{R}_n$ then the optimal weight vector can be found by differentiating equation (2.3.4) with respect to the weight vector and setting the gradient to zero. The optimal weight vector is given by

$$\mathbf{a}_{opt} = \mathbf{R}_n^{-1} \mathbf{p}_n \tag{2.3.5}$$

In a stationary environment $\mathbf{R}_n = \mathbf{R}$ and $\mathbf{p}_n = \mathbf{p}$ and $\mathbf{a}_{opt} = \mathbf{R}^{-1}\mathbf{p}$. As stated earlier, the Least Mean Squares algorithm uses an estimate of the gradient of the error at each update of the tap weights. The tap weight $\mathbf{a}_n$ is updated as

$$\mathbf{a}_{n+1} = \mathbf{a}_n - \mu \hat{\nabla}_{\mathbf{a}} E\left[z_n^2\right] \tag{2.3.6}$$

Where $\hat{\nabla} E[z_n^2]$ is the estimate of the true gradient. Approximating $E[z_n^2]$ as $z_n^2$ allows one to compute the gradient estimate as follows

$$\begin{aligned} \hat{\nabla}_{\mathbf{a}} E\left[z_n^2\right] &= \nabla_{\mathbf{a}} z_n^2 \\ &= 2 z_n \nabla_{\mathbf{a}} z_n \end{aligned} \tag{2.3.7}$$

With $z_n = p_n - \mathbf{a}_n^T \mathbf{r}_n$, equation (2.3.7) becomes

$$\hat{\nabla}_{\mathbf{a}} E\left[z_n^2\right] = -2 z_n \mathbf{r}_n \tag{2.3.8}$$

Putting equation (2.3.8) into (2.3.6) gives the LMS algorithm

$$\mathbf{a}_{n+1} = \mathbf{a}_n + 2\mu z_n \mathbf{r}_n \tag{2.3.9}$$

Now, our weight vector $\mathbf{a}_n$, is a function of random variables so is itself a random variable. It can be shown [8] that $\mathbf{a}_n$ is an asymptotically unbiased estimate of $\mathbf{a}_{opt}$.

The condition for equation (2.3.9) to converge is that [8]

$$0 < \mu < \frac{1}{\lambda_{max}} \tag{2.3.10}$$

where $\lambda_{max}$ equals the maximum eigenvalue of $\mathbf{R}$. In practice, one does not typically know the value of the maximum eigenvalue of $\mathbf{R}$. To set the value of $\mu$, we note that the trace of $\mathbf{R}$ equals the total input power to the tap weights and because $\mathbf{R}$ is positive definite, $tr[\mathbf{R}] > \lambda_{max}$. Therefore, a sufficient condition for convergence of (2.3.9) is

$$0 < \mu < \frac{1}{tr[\mathbf{R}]} \tag{2.3.11}$$

The variance of $\mathbf{a}_n$ as $n \to \infty$ is derived by Widrow [2] for the case when $z_n$ and $\mathbf{r}_n$ are Gaussian. In this case, the variance of $\mathbf{a}_n$ is

$$\lim_{n \to \infty} \mathrm{Var}\left[\mathbf{a}_n\right] = \left(I - \mu\Lambda\right)^{-1} \mu \varsigma_{min} I \tag{2.3.12}$$

where $I$ equals the $L \times L$ identity matrix, $\varsigma_{min}$ equals the minimum value of $E[z_n^2]$, and $\Lambda$ equals a diagonal matrix of the eigenvalues of $\mathbf{R}$. If we assume that $\mu$ is small such that $\mu << I$, then equation (2.3.12) simplifies further to

$$\lim_{n \to \infty} \mathrm{Var}\left[\mathbf{a}_n\right] = \mu \varsigma_{min} I \tag{2.3.13}$$

Equation (2.3.13) tells us that the variance of our final solution is directly proportional to $\mu$. Unfortunately, convergence of the LMS algorithm is inversely proportional to $\mu$. Hence, there is a tradeoff between the convergence rate and the variance in the estimate of the optimal weight vector.

Widrow [8] introduced a term called misadjustment which is defined as the ratio of the average excess mean square error to the minimum mean square error. Mathematically, the misadjustment $M$, is

$$M = \frac{\varsigma - \varsigma_{min}}{\varsigma_{min}} \tag{2.3.14}$$

From [8], it can show that $M = \mu tr[\mathbf{R}]$. In this report, typical values for $M$ were between 2.5% and 10%. Empirically, these values for $M$ were found to be close enough to the optimal solution and to have a reasonable convergence rate for the experiments conducted in this study.

In closing, the LMS algorithm offers an advantage in its simplicity and efficiency in recursively estimating the optimal weight vector. A disadvantage with the algorithm is that its convergence rate is governed by the inverse of the smallest eigenvalue of $R$. In non-stationary environments, the LMS algorithm may not be able to track the

optimal weight vector. The next section discusses the recursive least squares algorithm which minimizes a slightly different error criterion but which has excellent convergence properties at the expense of greater complexity and computation.

## 2.4 Summary of the Recursive Least Squares Algorithm

In the previous section, we considered a method which asymptotically approaches the optimal weight vector for minimizing $E[z_n^2]$. Now, let us consider a different error criterion. Let our error criterion be

$$\varepsilon(\mathbf{a}) = \sum_{i=0}^{N-1} \alpha^i z_{n-i}^2 \tag{2.4.1}$$

where $0 < \alpha \leq 1$. The variable $\alpha$ represents an exponential weighting of past samples. If the environment is stationary so that all past data is equally important, then one would select $\alpha = 1$ in equation (2.4.1).

Noting that $z_n = \mathbf{r}_n^T \mathbf{a} - p_n$, we can rewrite $\varepsilon(\mathbf{a})$ as

$$\varepsilon(\mathbf{a}) = \left[\mathbf{R}_n \mathbf{a} - \mathbf{p}_n\right]^T \mathbf{A}_n \left[\mathbf{R}_n \mathbf{a} - \mathbf{p}_n\right] \tag{2.4.2}$$

where

$$\mathbf{R}_n = \begin{pmatrix} \mathbf{r}_{n-(N-1)}^T \\ \mathbf{r}_{n-(N-2)}^T \\ \vdots \\ \mathbf{r}_n^T \end{pmatrix} \qquad \mathbf{p}_n = \begin{pmatrix} p_{n-(N-1)} \\ p_{n-(N-2)} \\ \vdots \\ p_n \end{pmatrix}$$

$$\mathbf{A}_n = \begin{pmatrix} \alpha^{N-1} & 0 & \dots & 0 & 0 \\ 0 & \alpha^{N-2} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \alpha & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

and $p_n$ is defined as the value of the primary input at time $n$. The term $\mathbf{r}_n$ is defined as the vector of inputs to the individual taps in the adaptive filter at time $n$. The parameter $\mathbf{a}$ is defined as the vector of tap weights.

To minimise (2.4.2), we set the derivative of $\varepsilon(\mathbf{a})$ with respect to $\mathbf{a}$ equal to zero. The least squares solution for $\mathbf{a}$ is

$$\mathbf{a} = \left[\mathbf{R}_n^T \mathbf{A}_n \mathbf{R}_n\right]^{-1} \mathbf{R}_n^T \mathbf{A}_n \mathbf{p}_n \tag{2.4.3}$$

The idea behind the recursive least squares method is to update (2.4.3) in an efficient manner when a new data point becomes available. Following Monzingo and Miller [9], we can partition the update signals $\mathbf{R}_{n+1}$, and $\mathbf{p}_{n+1}$ as follows

$$\mathbf{R}_{n+1} = \begin{pmatrix} \mathbf{R}_n \\ \cdots \\ \mathbf{r}_{n+1}^T \end{pmatrix}$$

$$\mathbf{p}_{n+1} = \begin{pmatrix} \mathbf{p}_n \\ \cdots \\ p_{n+1} \end{pmatrix}$$

Now define

$$\mathbf{P}_n = \left[\mathbf{R}_n^T \mathbf{A}_n \mathbf{R}_n\right]^{-1}$$

$$\mathbf{P}_{n+1} = \left[\mathbf{R}_{n+1}^T \mathbf{A}_{n+1} \mathbf{R}_{n+1}\right]^{-1} \tag{2.4.6}$$

Letting $\gamma_{n+1}$ be defined as $[\alpha + \mathbf{r}_{n+1}^T \mathbf{P}_n \mathbf{r}_{n+1}]^{-1}$, the update for $\mathbf{a}_{n+1}$ can be computed as [9]

$$\mathbf{a}_{n+1} = \mathbf{a}_n + \gamma_{n+1} \mathbf{P}_n \mathbf{r}_{n+1} \left[p_{n+1} - \mathbf{a}_n^T \mathbf{r}_{n+1}\right] \tag{2.4.5}$$

$$\mathbf{P}_{n+1} = \frac{1}{\alpha}\left[\mathbf{P}_n - \gamma_{n+1} \mathbf{P}_n \mathbf{r}_{n+1} \mathbf{r}_{n+1}^T \mathbf{P}_n\right] \tag{2.4.6}$$

Equation (2.4.5) and (2.4.6) are the update equations for the Recursive Least Squares algorithm.

The initialization of the recursion given in (2.4.5) and in (2.4.6) can be done two ways. The first method is to take the first $n$ points and solve for $\mathbf{a}_n$ and $\mathbf{P}_n$ directly from

$$\mathbf{P}_n = \left[\mathbf{R}_n^T \mathbf{A}_n \mathbf{R}_n\right]^{-1}$$

$$\mathbf{a}_n = \mathbf{P}_n \mathbf{R}_n^T \mathbf{A}_n \mathbf{p}_n \tag{2.4.7}$$

The second method is to set $\mathbf{a}_0$ to some arbitrary value and set $\mathbf{P}_0 = \beta I$ where $\beta$ is a large positive scalar and $I$ is the identity matrix. Now, using (2.4.6), $\mathbf{P}_n$ is [13]

$$\mathbf{P}_n = \left[\mathbf{P}_0^{-1} + \mathbf{R}_n^T \mathbf{A}_n \mathbf{R}_n\right]^{-1} \tag{2.4.8}$$

and $\mathbf{a}_n$ is

$$\mathbf{a}_n = \mathbf{P}_n \left[\mathbf{R}_n^T \mathbf{A}_n \mathbf{p}_n + \mathbf{P}_0^{-1} \mathbf{a}_0\right] \tag{2.4.9}$$

To make (2.4.8) agree with (2.4.7), we need to force $\mathbf{P}_0^{-1}$ to zero. This can be done by letting $\beta \to \infty$ since

$$\lim_{\beta \to \infty} \mathbf{P}_0^{-1} = \lim_{\beta \to \infty} \frac{1}{\beta} I = \mathbf{0} \tag{2.4.10}$$

Empirically, with $\beta = 10^{10}$, the difference in the values of the coefficients from using (2.4.8) versus using (2.4.7) for the updates was found to be less than one part in $10^{-5}$.

The value of $\alpha$ chosen reflects on how rapidly the environment is changing. Let us define the effective distance into the past data as $1/(1 - \alpha)$. In this study, $\alpha$ was

chosen such that the effective length is ten times the number of taps. For example, using 100 tap weights, $\alpha$ is set to 0.999. The RLS algorithm converges to $\mathbf{a}_{opt}$ within a few multiples of $L$ points, where $L$ equals the number of tap weights. The disadvantage of the RLS algorithm is that updates of $\mathbf{a}_n$ require $O(L^2)$ operations. For comparison, the LMS method requires $O(L)$ operations per update.

### 2.5 Conclusion

This chapter has presented a summary of the adaptive noise cancelling algorithm. Also covered are two common adaptive filter algorithms, the LMS and the RLS algorithms. It should be noted that there are exact least squares methods that minimize (2.4.1) and only require $O(L)$ computations per update point [10,14]. This author has found empirically that these efficient least squares methods are numerically sensitive to their initial starting conditions. Because of the numerical stability and ease of initialization of the RLS method, it was chosen as a representative exact least squares algorithm to be compared with the LMS algorithm.

# CHAPTER 3

# NEW APPLICATIONS OF ADAPTIVE NOISE CANCELLATION

## 3.1 Introduction

In the previous chapter, the application of ANC to situations where the reference input contained no speech components was discussed. Boll and Pulsipher [6] created a laboratory environment to satisfy this requirement and showed that a significant reduction of background noise could be achieved using the ANC approach. In this chapter, a method of applying ANC to a more realistic environment is developed. Specifically, we consider the case when there is a non-ideal acoustic barrier between the noise source and the desired signal. In this case the primary and reference microphones can be placed on opposite sides of the acoustic barrier. An environment which exhibits these characteristics is the cockpit of a fighter aircraft in which the pilot wears an oxygen facemask. The application of ANC to this environment will be discussed more fully in Chapter 4.

If the acoustic barrier is not totally effective then both the primary and reference microphones will contain speech and noise components. The model of the system is shown in Figure 3.1. Here we have assumed that there is a transfer function relating the noise in the reference input to the noise in the primary input. Conversely, there exists a transfer function that maps the speech in the primary input to the speech in the reference input. Updating the adaptive filter continuously can produce significant

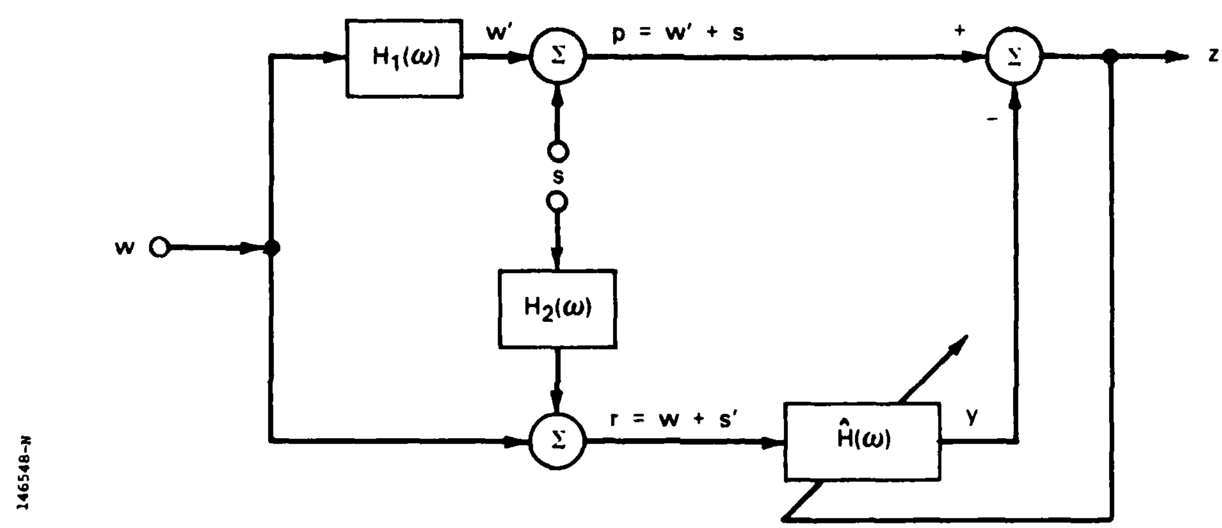



**Figure 3.1. Block diagram of adaptive noise cancellation with speech in the reference input**

speech distortion since the filter $\hat{H}(\omega)$ will attempt to reproduce a signal containing speech as well as noise. Unfortunately, this distortion is inversely proportional to the noise level. In the limit when the noise is zero, there is just speech present and it is completely cancelled by the system. Clearly, in environments such as in fighter aircraft where the ambient noise field can vary in intensity by 15-20 dB, adaptive noise cancellation as was previously used is not a viable method.

### 3.2 Modified Adaptive Noise Cancellation

To avoid distorting the speech, a system was developed in which the adaptive filter coefficients were updated only during intervals when there was no speech present. It is implicitly assumed in this approach that the transfer function $H_1(\omega)$ in Figure 3.1 does *not change significantly* when speech is present. During silence, $s = 0$ and the adaptive filter $\hat{H}(\omega)$ will try to approximate $H_1(\omega)$. When speech is present, the adaptation of the filter is halted by setting the adaptation constant $\mu$ in equation (2.3.9) to zero. Halting adaptation of the RLS method is done by not adding the correction term $\gamma_{n+1}\mathbf{P}_n\mathbf{r}_{n+1}[p_{n+1} - \mathbf{a}_n^T\mathbf{r}_{n+1}]$ in equation (2.4.5) to $\mathbf{a}_n$. More details on the speech detection algorithm used in this report are discussed in Chapter 4.

Using Figure 3.1, the primary input to our system contains speech $s$ plus noise $w'$. The reference input also contains speech $s'$ plus noise $w$. Now, $w' << w$ in energy due to the acoustic barrier. For example, with a pilot wearing a facemask, the rubber of the mask attenuates the ambient noise such that the sound pressure level (SPL) due to the noise inside the mask is lower than the SPL due to the noise outside of the mask. Also,

$s' << s$ in energy due to the acoustic barrier. In our example, the oxygen facemask attenuates the pilot's speech such that the SPL due to the speech inside of the mask is much higher than the SPL due to the speech outside of the mask.

Now, let us consider the case when no speech is present. For this case, the primary and reference inputs are given by

$$p = w'$$
$$r = w \tag{3.2.1}$$

In this case the adaptive filter $\hat{h}$ will try to minimize the output energy. With no speech present, the output $z$ is

$$\begin{aligned} z &= p - y \\ &= w' - w * \hat{h} \\ &= w * h_1 - w * \hat{h} \end{aligned} \tag{3.2.2}$$

where $*$ represents the convolution operator. With the energy of $z$ minimized, $w*\hat{h} \approx w'$ or equivalently, $\hat{h} \approx h_1$. What the above implies is that the adaptive filter $\hat{h}$ must attenuate the noise signal $w$ to make it match $w'$.

When speech occurs, the update of $\hat{h}$ is stopped. During speech, the output $z$ is given by

$$\begin{aligned} z &= p - y \\ &= s + w' - w * \hat{h} - s' * \hat{h} \end{aligned} \tag{3.2.3}$$

If we assume that $h_1$ does not change significantly when speech occurs, then $w' - w * \hat{h}$ will still be small compared with $s$. Now, the output is approximately given by

$$z \approx s - s' * \hat{h} \tag{3.2.4}$$

As we have stated earlier, $s' << s$ in energy. The effect of $\hat{h}$ on $s'$ is to attenuate $s'$ further such that $s' * \hat{h}$ is much smaller in amplitude than $s$ and $z \approx s$.

In the frequency domain, the Fourier transform of the output $Z(\omega)$ when speech is present is given by

$$\begin{aligned} Z(\omega) &= S(\omega) - S(\omega)H_2(\omega)\hat{H}(\omega) + W(\omega)H_1(\omega) - W(\omega)\hat{H}(\omega) \\ &= S(\omega)[1 - H_2(\omega)\hat{H}(\omega)] + W(\omega)[H_1(\omega) - \hat{H}(\omega)] \end{aligned} \tag{3.2.5}$$

Again, assuming $\hat{H}(\omega) \approx H_1(\omega)$ such the $W(\omega)[H_1(\omega) - \hat{H}(\omega)]$ is small compared with $S(\omega)[1 - H_2(\omega)\hat{H}(\omega)]$, the Fourier transform of the output $Z(\omega)$ is approximately

$$Z(\omega) \approx S(\omega)[1 - H_2(\omega)H_1(\omega)] \tag{3.2.6}$$

Intuitively, we want the magnitude of $H_2(\omega)H_1(\omega)$ to be much less than one for all frequencies in order for $Z(\omega)$ to approximately equal $S(\omega)$. This provides the basis behind the approach of using ANC in an environment where the combined attenuation of $H_1(\omega)$ and $H_2(\omega)$ is much less than one for all frequencies. Chapter 4 discusses an application of ANC in a fighter cockpit environment to test the above approach to noise reduction.

## 3.3 Conclusion

This chapter discussed a modification of the adaptive noise cancelling algorithm. The modification was to adapt the system during the silence intervals in the speech. Two conditions need to be met in order for this system to work. The first is that $H_1(\omega)$ in Figure 3.1 does not change appreciably between silence intervals in the speech. The second is that the magnitude of $H_1(\omega)H_2(\omega)$ is much less than one for all frequencies. This last condition assures that the leakage of speech from the reference input does not degrade the final output speech. The next chapter discusses an application of the ideas presented in this chapter to the fighter aircraft environment.

# CHAPTER 4

# APPLICATION OF ANC TO THE FIGHTER COCKPIT ENVIRONMENT

## 4.1 Introduction

The condition of $|H_1(\omega)H_2(\omega)| << 1$ given in the previous section occurs frequently in practice. For example, the oxygen facemask worn by a pilot in a military fighter aircraft provides an acoustic barrier which attenuates signals by about 10 dB. This attenuation applies to both $H_1(\omega)$ (ambient noise) and $H_2(\omega)$ (speech) so that the combined attenuation due to $|H_1(\omega)H_2(\omega)|$ is approximately 20 dB. Since appropriate in-flight data from an actual fighter aircraft was unavailable, the environment was partially simulated in a laboratory in order to evaluate the modified ANC method proposed in Chapter 3.

## 4.2 Experimental Set-up

Experiments were conducted in a small (10′ by 10′) partially sound-proofed room. An Allison 4 loudspeaker mounted on one wall was used to create the ambient noise field. A subject was seated near the loudspeaker and wore a regulation Air Force oxygen facemask. Placed inside the mask was a pressure gradient (noise cancelling) microphone identical to the ones used by Air Force pilots. Attached to the outside of the mask was an omnidirectional electret microphone. The electret microphone was

used because of its flat response from 100 to 4000 Hz and because of its omnidirectional behaviour. The outside microphone was positioned to be as close as possible to the inside microphone. The smallest separation was 3 cm. The reason for placing the two microphones as close as possible is twofold. The first is that smaller filter lengths for the adaptive filter can be used. The second is the potentially improved performance in a multi-noise source environment. In a complex noise environment with many different sources distributed over a region, the noise becomes uncorrelated between two points in space as we move the points further apart. Intuitively, if the microphones were separated by zero length they would both record the same signal and would be perfectly correlated with each other. As we separate the two microphones, they contain less and less common information. Hence, there is a need to place the two microphones as close as possible to each other.

The microphone outputs were connected separately to a two-channel tape recorder. Due to the attenuation of the mask and the power limitation of the loudspeaker, recordings from the primary microphone inside the mask contained very little noise. As a result it was necessary to make recordings of speech alone and noise alone and combine the two signals digitally prior to analysis. Although the technique produced an artificial signal, it had the benefit of allowing a range of signal-to-noise levels to be evaluated without requiring re-recording. In addition, the availability of the noise-free speech signal permitted accurate computation of signal-to-noise ratios.

Another difficulty encountered during the recording sessions was that it was not

possible to determine empirically the actual attenuation function of the facemask worn by the subject because the gains of the individual tape recorder channels were unknown. Assuming the frequency response of each of the two tape recorder amplifiers to be approximately equal, the one unknown parameter was the relative gain between the primary and reference inputs. This parameter was assumed to be the attenuation function of the facemask. The attenuation was estimated by referring to the results of a study performed for M.I.T. Lincoln Laboratory by researchers at Bolt, Beranek, and Newman [15], who found that the amount of attenuation caused by an Air Force oxygen facemask varied from 9 to 11 dB. An 8 dB mask attenuation was selected as a conservative estimate for the current study. As a result, the energy of the speech signal at the reference microphone was adjusted to be 8 dB lower than the energy of the speech signal at the primary. Likewise, the energy of the noise signal at the primary microphone was adjusted to be 8 dB lower than the energy of the noise signal at the reference. As will be discussed in Chapter 5, with 8 dB set as the mask attenuation factor, the attenuation associated with the function $H_1(\omega)H_2(\omega)$ is sufficient to allow substantial noise cancellation with minimal distortion of the speech signal.

In Chapter 3 it was shown that by constraining the algorithm to adapt only during silence it is implicitly assumed that $H_1(\omega)$ does not change significantly when speech occurs. For the sentences used in this report, the maximum time of speech activity was less than 600 milliseconds. This means that adaptation of $\hat{H}(\omega)$ was stopped for no more than six tenths of a second before a gap in the speech occurred. Since $H_1(\omega)$

relates to the position of the speaker with respect to the noise source, it can be assumed that $H_1(\omega)$ does not change significantly during this interval. The results presented in Chapter 5 will verify this assumption.

### 4.3 Detection of Speech

As was previously discussed, the noise estimation filter was allowed to adapt only in the absence of speech. In general, accurate detection of speech over a range of background noise levels is a very difficult problem. In this report, an estimate of the energy of the primary signal was used to determine if speech was present. It is known [16] that using the energy of the signal is not a reliable way of detecting low energy portions of the speech such as weak fricatives at the beginning or at the end of an utterance. Fortunately, adapting the system when there is low energy speech present does not noticeably degrade the speech. This is because when the energy of the speech is below the noise level, the noise is the dominant component of the signal and the adaptive filter is not significantly influenced by the low level speech.

Details of the speech detection algorithm are as follows. The energy of the primary signal at time $n$ is computed as the sum of the square of $p_n$ five milliseconds before and after time $n$. In other words

$$E(n) = \sum_{i=-50}^{50} p_{n+i}^2,$$

where the sampling rate of the system is 10,000. A representative plot of $E(n)$ versus time is shown in Figure 4.1. Superimposed on the graph is the lower and upper

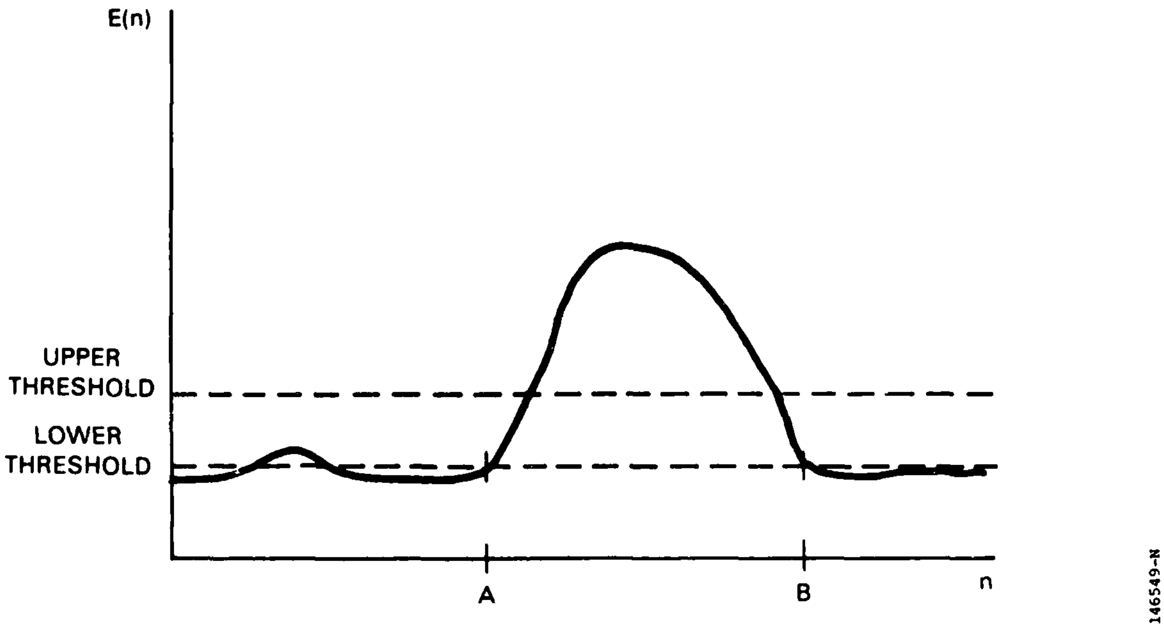



**Figure 4.1. Representative plot of $E(n)$ versus time and its use in determining when speech occurs**

thresholds. For the examples used in this report, the lower threshold was set to 1.2 times the estimate of the background noise power. The upper threshold was set to 1.5 times the estimate of the noise power. The beginning and ending points of speech are chosen as follows. When $E(n)$ is greater than the upper threshold, we say that speech is present. The beginning point of speech is arbitrarily chosen to be the point when $E(n)$ last crosses the lower threshold before it crosses the upper threshold. Looking at Figure 4.1, we see that the point A would mark the beginning of speech. The ending point of speech is chosen to be the point when $E(n)$ first crosses the lower threshold after it crosses the upper threshold. Looking at Figure 4.1, we see that the point B would mark the ending of speech.

To allow a gradual change in noise power, the estimate of the noise power is updated by low pass filtering $E(n)$ when no speech is present. The output of the low pass filter is used as the estimate of the noise power. The initial estimate of the noise power is computed by averaging the energy of the first second of recorded data which is assumed to be just noise.

### 4.4 Conclusion

This chapter covered the experimental set-up to test the modified adaptive noise cancellation algorithm discussed in Chapter 3. The oxygen facemask worn by the pilot provides a moderate amount of attenuation of the speech and noise signals. Direct application of the adaptive noise cancelling algorithm produces noticeable degradation of the speech signal. This degradation is most severe when the ambient noise level is

low. In Chapter 5, the results of using the modified ANC method on the experimental data will be presented and discussed.

# CHAPTER 5

# EXPERIMENTAL RESULTS AND DISCUSSION

## 5.1 Introduction

In this chapter we present the results of using the modified adaptive noise cancelling algorithm on the data described in Chapter 4. As was noted in Chapter 2, there exists many ways to design the adaptive filter embedded in the ANC algorithm. This report will use the Least Mean Squares (LMS) and the Recursive Least Squares (RLS) algorithms to implement the adaptive filter.

## 5.2 Experimental Results

With the set-up described in Chapter 4, the test subject spoke the sentence *'He scanned through the law books.'* To allow the adaptive filter algorithms time to initially adapt, the first one and half seconds of data contain only noise. The last two and half seconds of data contain the speech plus noise. Using the LMS and RLS algorithms and adapting during silence as explained in the previous chapter, the signal-to-noise ratio (SNR) improvement versus time provided by the ANC system for various filter lengths is given in Figures 5.1a,b. For the examples in Figures 5.1a,b, the initial SNR is 3 dB.

Since the noise and undegraded speech are available, the SNR improvement used in this report is computed as follows. For both the input and output, the ratio of the energy in the speech to the energy in the noise is computed over a 100 millisecond

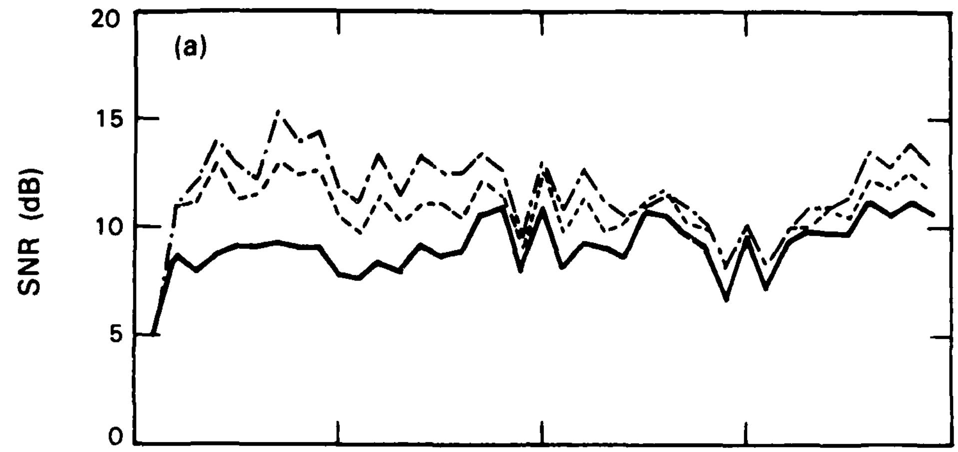



**Figure 5.1a.** Improvement in signal-to-noise ratio versus time for $L = 25$, 50, and 100 using the LMS algorithm with $M$ adjusted to equal 5% for each case. Initial SNR equals 3 dB.

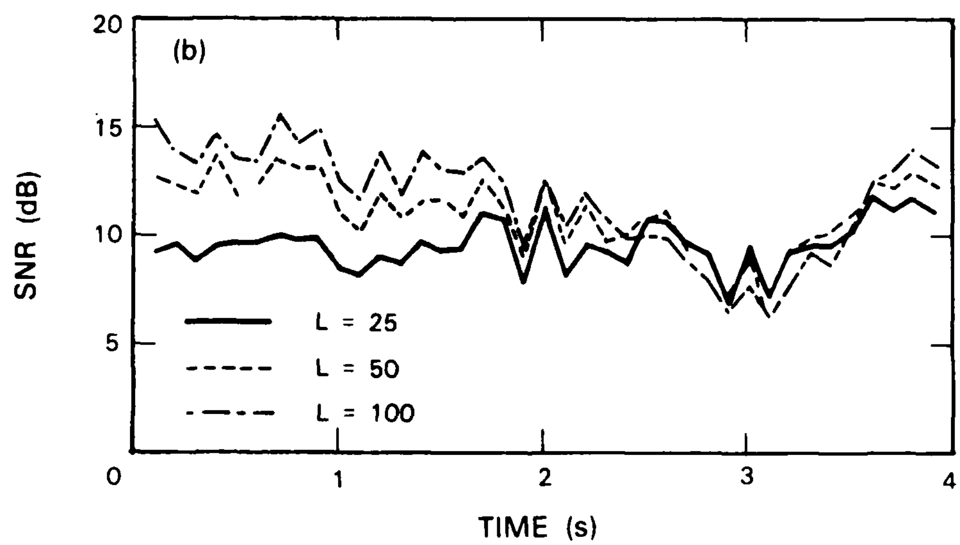



**Figure 5.1b.** Improvement in signal-to-noise ratio versus time for $L = 25$, 50, and 100 using the RLS algorithm with $\alpha$ adjusted such that $1/(1 - \alpha)$ equals $10L$. Initial SNR equals 3 dB.

frame. The log ratio of the output SNR to the input SNR is multiplied by 10 to get an improvement in SNR in dB. During the interval where only silence occurs, the SNR improvement in dB is the log ratio of the input noise energy to the output noise energy multiplied by 10. The initial SNR is computed as follows. The energy of the original clean speech signal is normalized to have unit energy. The energy of the noise signal is adjusted such that ten times the log ratio of the speech energy to the noise energy is 3, 6, or 10 dB. Also, as discussed in Chapter 4, the mask attenuation is set at 8 dB. Therefore, the signal-to-noise ratio at the reference input is 16 dB below the SNR at the primary input.

Another observation to be made is that our system must be causal. For example, if the noise reached the primary microphone first and then after a slight delay, reached the reference microphone then our system would be non-causal. To force causality, a small delay is placed in the primary channel. For a filter length of 25 taps, the delay is 1.2 milliseconds. For the filter lengths of 50 and 100 taps, the delay is 2.5 milliseconds. The exact value of the delay is not critical. The delay only needs to be long enough to make the system causal.

From Figure 5.1a, the average improvement in SNR is 9.3 for $L = 25$, 11.0 for $L = 50$, and 11.9 for $L = 100$ when using the LMS algorithm. Using the RLS algorithm, the average improvement in SNR is 9.7 for $L = 25$, 11.2 for $L = 50$, and 11.7 for $L = 100$. We see that for $L = 25$ and 50, the RLS method performs better than the LMS method. The reason for this increase in performance is due to the compromise between

the convergence rate and the variance of the filter estimate in the LMS algorithm. For $L = 100$, the RLS method actually performs worse than the LMS method when speech is present. The reason for this behavior is that the RLS method uses past data to determine the update of the filter coefficients. As was stated in Chapter 2, the amount of past data used is proportional to the number of tap weights. When adaptation resumes, noise corrupted by speech is still being used to update $\hat{H}(\omega)$ for larger values of $L$. The net result is that the filter coefficients derived are not correct. The effect is that for $L = 100$, there is a slight reverberation in the processed speech. For $L = 25$ and 50, the above effects are much smaller and no noticeable reverberation is heard in the processed speech. Another point to be made is that above 50 taps, there is a marginal increase in performance for a given increase in the filter length.

The results obtained show that a filter length of 50 taps is all that is needed to give approximately 11.0 dB improvement in SNR. Furthermore, the amount of improvement in SNR is approximately independent of the input SNR. From Figures 5.2a,b, the average amount of improvement in SNR is 11.0 dB for an input SNR of 3 dB, 11.0 dB for an input SNR of 6 dB, and 10.9 dB for an input SNR of 10 dB for the LMS method. For the RLS method, the average amount of improvement in SNR is 11.2 dB for an input SNR of 3 dB, 11.1 dB for an input SNR of 6 dB, and 11.1 dB for an input SNR of 10 dB. The differences in improvement in SNR across the different input SNR's are due to the speech detection algorithm.

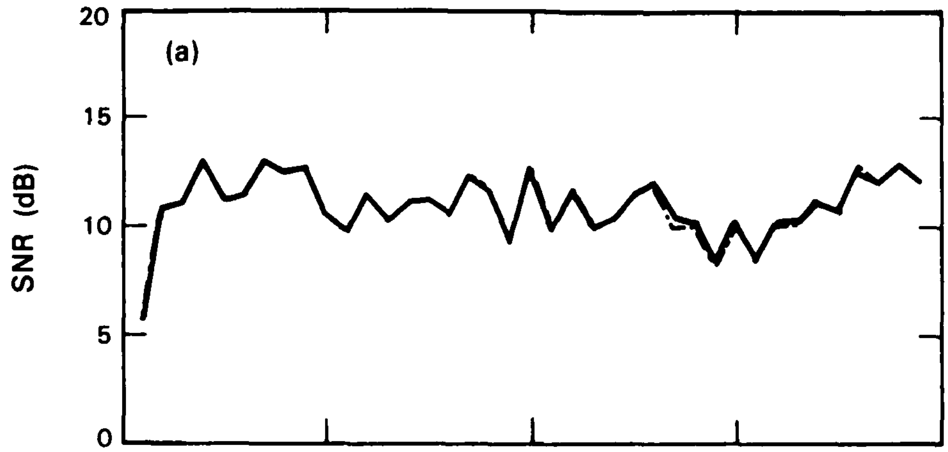
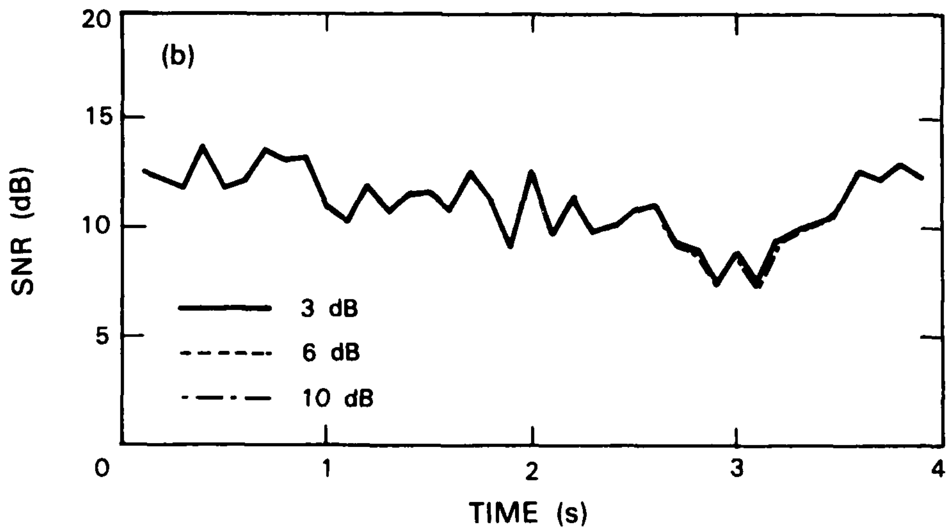



**Figure 5.2a.** Improvement in signal-to-noise ratio versus time for an initial SNR of 3, 6, and 10 dB. LMS algorithm used with $L = 50$, and $M = 5\%$.




146551-N

**Figure 5.2b.** Improvement in signal-to-noise ratio versus time for an initial SNR of 3, 6, and 10 dB. RLS algorithm used with $L = 50$, and $\alpha = 0.998$.

## 5.3 Discussion

To compare the rate of convergence of the LMS and RLS algorithms, let us plot the output signal for the first 200 milliseconds as shown in Figures 5.3a,b. As was mentioned earlier, the input data initially contains only noise and no speech. The number of tap weights was set to 100 and the initial SNR was 3 dB. As the rate of convergence of our adaptive filters is inversely proportional to the filter length, the convergence rate for $L = 25$ and $L = 50$ will be greater than that for $L = 100$. For the LMS algorithm, $\mu$ in equation (2.3.11) was set such that the misadjustment was 5%. For the RLS algorithm, $\alpha = 0.999$. It should be noted that for both the LMS and RLS methods, the initial value for the tap weights was zero. Also, the output of both systems was normalized by the same constant. In Figure 5.3a, we see that the LMS algorithm requires approximately 120 milliseconds before the energy of the output signal reaches a steady state value. For comparison, in Figure 5.3b, we see that the RLS method required only 20 milliseconds to converge. It is interesting to note that for the first 10 milliseconds (or first 100 points), the output using the RLS method is zero. This is because our filter length is 100 taps and for $n < 100$, we have enough degrees of freedom to set the output error exactly to zero.

Clearly, in environments that may experience sudden changes, an adaptive filter that can rapidly adapt should be used. Also, if the spread of the eigenvalues of $\mathbf{R}$ (see Section 2.3) is large, then the LMS algorithm may not converge quickly enough. In these experiments, sudden changes in the environment did not occur. Also, the

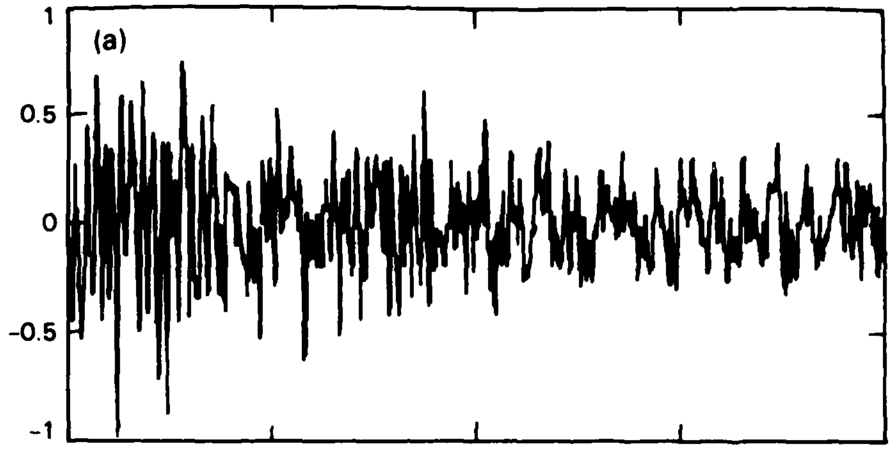



**Figure 5.3a. Normalised output of the ANC system using the LMS algorithm. Initial SNR = 3 dB. $L$ = 100, and $M$ = 5%.**

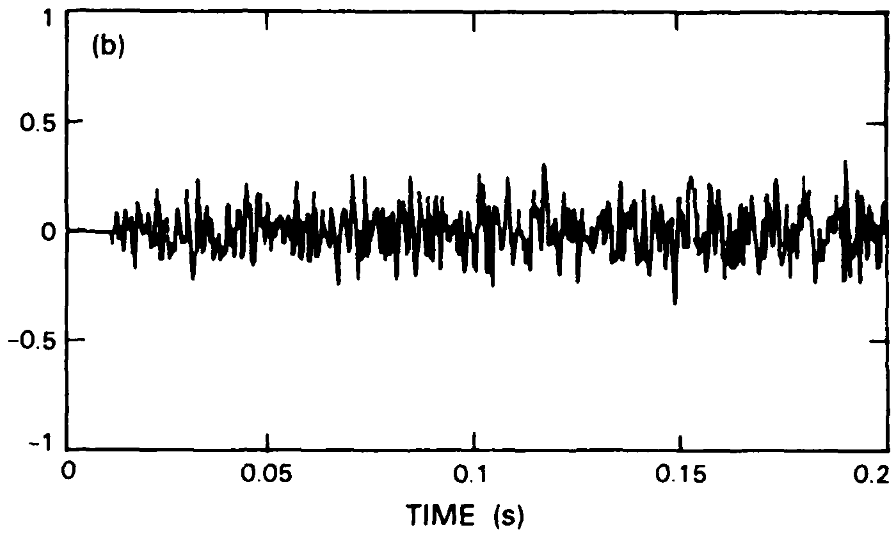



**Figure 5.3b. Normalised output of the ANC system using the RLS algorithm. Initial SNR = 3 dB. $L$ = 100, and $\alpha = 0.999$.**

spectrum of the noise played through the loudspeaker was flat. Hence, there was not a large spread in the eigenvalues of $\mathbf{R}$, and overall, the convergence rate of the adaptive filter was not a critical parameter.

Little speech distortion was heard in the processed speech during informal listening tests. This is primarily due to constraining the algorithm to adapt only during silence. The slight distortion heard in the processed speech can be attributed to the presence of a speech signal component at the reference microphone.

Even though the results in these experiments are quite encouraging, they cannot be readily applied to the true fighter cockpit environment. Whereas the noise component in this study was derived from a single loudspeaker, the noise inside the cockpit results from a complicated interaction of many sources and may vary both spatially and temporally. Since the successful application of the ANC method depends directly on the degree to which the differences between the signals at the primary and reference microphones can be modeled by a linear filter, it is intuitively obvious that the greater the separation between the two microphones in a multiple source environment the lower the likelihood that the use of a single adaptive filter will suffice to cancel the noise without introducing distortion. However, a close placement of the microphones on the facemask may be adequate even in a sound field as complex as that of a fighter cockpit. Validation of this intuition must await further experimentation in either a more accurately simulated cockpit environment or, preferably, on board an actual aircraft.

## CHAPTER 6

## CONCLUSION AND PROPOSED RESEARCH

In this report it has been shown that the modified ANC algorithm can provide substantial noise cancellation under conditions which more closely approximate those found in practice than the conditions used by previous researchers [6]. The fundamental requirement is that there be sufficient attenuation of the speech signal between the primary and reference inputs. Results showed that this requirement is met for the case of the oxygen facemask worn by fighter aircraft pilots and that the facemask allows placement of the reference microphone close to the primary microphone. The resulting short filter length begins to make it practical to consider real time implementation of the ANC algorithm.

The experiments reported in this study were conducted using a single noise source in a laboratory sound chamber. It is of considerable interest to determine whether the proposed technique is effective in a multiple or distributed source environment more nearly like that of a fighter cockpit. As of yet, no theory exits that predicts the performance of the ANC algorithm in a general environment. For instance, if one knew that the ambient noise was isotropic with a certain spectral covariance structure, then it would be useful to be able to predict the improvement in SNR for a given microphone configuration.

Because of the limitations discussed in Chapter 4, the noise and speech were recorded separately in the experiments conducted. In addition, laboratory-generated

white noise was used as the noise source rather than the noise that arises in an actual fighter aircraft. Even though it is anticipated that the results will generally hold when speech and noise are simultaneously recorded and the noise used is noise in a real cockpit environment, it would be useful to establish this point.

## ACKNOWLEDGMENTS

The author would like to thank the Digital Signal Processing Group at MIT and Group 24 at MIT Lincoln Laboratories for providing the facilities to carry out this research. It has been a pleasure to work with the members of both groups. Special thanks goes to Thrasyvoulos Pappas, and Professor Jae Lim of the Digital Signal Processing Group, and Elliot Singer, Tom Quatieri, and Doug Paul of Group 24 for providing many helpful discussions.

## REFERENCES


[1] J. S. Lim, ed., Speech Enhancement (Prentice-Hall, Englewood Cliffs, N. J., 1983).

[2] B. Widrow, et al, "Adaptive Noise Cancelling: Principles and Applications," *Proc. IEEE*, 63, pp. 1692–1716, December 1975.

[3] M. M. Sondhi, "An Adaptive Echo Canceller," *B.S.T.J.*, 66, pp. 497–511, March 1967.

[4] J. R. Treicher, "The Spectral Line Enhancer – The Concept, an Implementation, and an Application," *Ph.D dissertation*, Dept. of Elec. Eng., Stanford Univ., Stanford, Calif., June 1977.

[5] L. J. Griffiths, "An Adaptive Noise Cancelling Procedure for Multidimensional Systems," *Proc. Circuits and Systems Conf.*, Asilomer, Calif., Nov. 1976.

[6] S. F. Boll and D. C. Pulsipher, "Suppression of Acoustic Noise in Speech Using Two Microphone Adaptive Noise Cancellation," *IEEE Trans. Acoustics, Speech and Signal Proc.*, ASSP-28, pp. 752-753, (1980).

[7] B. Widrow and M. E. Hoff, "Adaptive Switching Circuits," in *IRE WESCON Conv. Rec.*, pt. 4, pp. 96–104, (1960).

[8] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," *Proc. IEEE*, 64, pp. 1151-1162, (1976).

[9] R. A. Monzingo and T. W. Miller, Introduction to Adaptive Arrays (Wiley and Sons, Inc., 1980).

[10] M. Morf and D. T. Lee, "Recursive Least Squares Ladder Forms for Fast Parameter Tracking," in *1978 IEEE Conf. Decision and Control*, San Diego, CA, pp. 1362–1367, Jan 10 12, 1979.

[11] B. Friedlander, "System Identification Techniques for Adaptive Signal Processing," *Circuits Systems and Signal Processing*, 1, pp. 3–41, (1982).

[12] L. J. Griffiths, "An Adaptive Lattice Structure for Noise-Cancelling Applications," *Proc. 1978 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Tulsa OK, pp. 87 90, April 1978.

[13] T. C. Hsia, System Identification (Lexington Books, Lexington, MA, 1977).

[14] D. Falconer and L. Ljung, "Application of Fast Kalman Estimation to Adaptive Equalization," *IEEE Trans. Comm.*, COMM-26, pp. 1439–1446, (1978).

[15] R. L. Miller, R. D. Bruce, F. N. Iacovino, and A. W. F. Huggins, "Simulation of Cockpit Noise Environment, in Four Tactical Aircraft for the Purpose of Testing Speech Intelligibility" (Bolt Beranek and Newman, Sept. 1981).

[16] L. R. Rabiner and M. R. Sambur, "An Algorithm for Determining the Endpoints of Isolated Utterances," *B.S.T.J.*, 54, pp. 297–315, (1975).

## REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

**1. REPORT NUMBER**
ESD-TR-84-040

**2. GOVT ACCESSION NO.**
AD-A149 223

**3. RECIPIENT'S CATALOG NUMBER**

**4. TITLE (and Subtitle)**
Speech Enhancement Using Multiple Microphones

**5. TYPE OF REPORT & PERIOD COVERED**
Technical Report

**6. PERFORMING ORG. REPORT NUMBER**
Technical Report 691

**7. AUTHOR(s)**
William A. Harrison

**8. CONTRACT OR GRANT NUMBER(s)**
F19628-85-C-0002

**9. PERFORMING ORGANIZATION NAME AND ADDRESS**
Lincoln Laboratory, M.I.T.
P.O. Box 73
Lexington, MA 02173-0073

**10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS**
Program Element Nos. 33401F and 63735F
Project No. 7820

**11. CONTROLLING OFFICE NAME AND ADDRESS**
Air Force Systems Command, USAF
Andrews AFB
Washington, DC 20331

**12. REPORT DATE**
15 November 1984

**13. NUMBER OF PAGES**
54

**14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)**
Electronic Systems Division
Hanscom AFB, MA 01731

**15. SECURITY CLASS. (of this report)**
Unclassified

**15a. DECLASSIFICATION DOWNGRADING SCHEDULE**

**16. DISTRIBUTION STATEMENT (of this Report)**
Approved for public release; distribution unlimited.

**17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)**

**18. SUPPLEMENTARY NOTES**
*This report is based on a thesis of the same title submitted to the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology in June 1984 in partial fulfillment for the degrees of Master of Science and Electrical Engineer.

**19. KEY WORDS (Continue on reverse side if necessary and identify by block number)**

| | |
|---|---|
| noise reduction | adaptive noise cancellation |
| least mean squares | recursive least squares |

**20. ABSTRACT (Continue on reverse side if necessary and identify by block number)**

A new application of Widrow's Adaptive Noise Cancelling (ANC) algorithm is presented. Specifically, the ambient environment is generalized to include the case where an acoustic barrier exists between the primary and reference microphones. By updating the coefficients of the noise estimation filter only during silence, it is shown that the ANC technique can provide substantial noise reduction with little speech distortion even when the acoustic barrier provides only moderate attenuation of acoustic signals. The use of the modified ANC method is evaluated using an oxygen facemask worn by fighter aircraft pilots. Experiments demonstrate that if a noise field is created using a single source, 11 dB signal-to-noise ratio improvement can be achieved by attaching the reference microphone to the exterior of the facemask. The length of the ANC filter required for this particular environment is only about 50 points long.

END

FILMED

2-85

DTIC